

10/581821

AP9 Rec'd PCT/PTO 02 JUN 2006

Attorney Docket No.: 082420-000400US

**PATENT APPLICATION**

**DEVICE CONTROL DEVICE, SPEECH RECOGNITION DEVICE,  
AGENT DEVICE, AND DEVICE CONTROL METHOD**

Inventor: Rika Koyama, a citizen of Japan, residing at  
Tokui Heights 205, 2-21, Tokuimachi 4-chome  
Nada-ku, Kobe-shi, Hyogo 657-0033, Japan

Assignee: Kabushikikaisha Kenwood  
2967-3, Ishikawa-machi, Hachioji-shi  
Tokyo 192-8525 Japan

Entity: Large

TOWNSEND and TOWNSEND and CREW LLP  
Two Embarcadero Center, Eighth Floor  
San Francisco, California 94111-3834  
Tel: 650-326-2400

# DEVICE CONTROL DEVICE, SPEECH RECOGNITION DEVICE, AGENT DEVICE, AND DEVICE CONTROL METHOD

## CROSS-REFERENCES TO RELATED APPLICATIONS

**[0001]** This present application claims priority to Japanese Application No. 2003-406730, filed December 5, 2003, commonly assigned and hereby of which is incorporated by reference in its entirety.

## BACKGROUND OF THE INVENTION

**[0002]** The present invention relates to a device control device, a speech recognition device, an agent device, and a device control method.

**[0003]** Recently, schemes of recognizing speeches using a speech recognition technology, and controlling electric devices or the like in response to the recognition result are used. For example, Patent Literature 1 (Unexamined Japanese Patent Application KOKAI Publication No. H8-339288) discloses a scheme of inputting a speech, identifying a word represented by the input speech, discriminating whether the identified word matches a predetermined keyword or not, and controlling an external device based on the discrimination result.

**[0004]** Patent Literature 1: Unexamined Japanese Patent Application KOKAI Publication No. H8-339288.

## BRIEF SUMMARY OF THE INVENTION

**[0005]** It is however difficult to completely recognize an instruction aurally uttered by a human being. Accordingly, there is a case where the scheme cannot adequately respond to an instruction uttered in the form of a language by a human being.

**[0006]** The present invention has been made in view of the above situation, and it is an object of the invention to provide a device control device, a speech recognition device, an agent device, and a device control method, which can control a device adequately in response to an instruction uttered in the form of a language by a human being.

**[0007]** To achieve the object, a device control device according to the first aspect of the invention comprises:

input information identification means which identifies input information to be input;

5 process-item data storing means which stores a plurality of process items for executing processes corresponding to recognized input information identified by the input information identification means;

transition-definition data storing means which stores plural pieces of transition definition data defining transition from one process item in the plurality of process items to  
10 another process item; and

update means which accesses an externally located source of a new process item or transition definition data, acquires the new process item or transition definition data, and updates an old process item stored in the process-item data storing means or old transition definition data stored in the transition-definition data storing means to the new process item  
15 or transition definition data, wherein

each of the transition definition data includes a condition corresponding to input information, and

a piece of transition definition data is selected based on comparison between the input information input and the conditions of the individual transition definition data, and a  
20 status is transitioned to a process item designated by the selected transition definition data.

**[0008]** The source may compress and supply the new process item or the transition definition data, and the update means may have means which decompresses the new process item or the transition definition data acquired from the source.

**[0009]** A speech recognition device according to the second aspect of the invention  
25 comprises:

speech signal identification means which identifies a speech signal to be input;

process-item data storing means which stores a plurality of process items for executing processes corresponding to a speech signal identified by the speech signal identification means;

30 transition-definition data storing means which stores plural pieces of transition definition data defining transition from one process item in the plurality of process items to another process item; and

update means which accesses an externally located source of a new process item or a transition definition data, acquires the new process item or transition definition data, and updates an old process item stored in the process-item data storing means or old transition definition data stored in the transition-definition data storing means to the new process item or transition definition data, wherein

each of the transition definition data includes a condition corresponding to input information, and

a piece of transition definition data is selected based on comparison between the speech signal input and the conditions of the individual transition definition data, and a status is transitioned to a process item designated by the selected transition definition data.

**[0010]** The source may compress and supply the new process item or the transition definition data, and

the acquisition means may have means which decompresses the new process item or the transition definition data acquired from the source.

**[0011]** An agent device according to the third aspect of the invention comprises:  
input information identification means which identifies input information to be input;

process-item data storing means which stores a plurality of process items for executing processes corresponding to input information identified by the input information identification means;

transition-definition data storing means which stores plural pieces of transition definition data each defining transition from one process item in the plurality of process items to another process item; and

update means which accesses an externally located source of a new process item or a transition definition data, acquires the new process item or transition definition data, and updates an old process item stored in the process-item data storing means or old transition definition data stored in the transition-definition data storing means to the new process item or transition definition data, wherein

each of the transition definition data includes a condition corresponding to input information, and

a piece of transition definition data is selected based on comparison between the input information input and the conditions of the individual transition definition data, and a status is transitioned to a process item designated by the selected transition definition data.

[0012] The source may compress and supply the new process item or the transition definition data, and the update means may have means which decompresses the new process item or the transition definition data acquired from the source.

[0013] A device control method according to the fourth aspect of the invention comprises:  
a process-item data storing step of storing a plurality of process items for executing processes corresponding to input information identified at an input information identification step;

a transition-definition data storing step of storing plural pieces of transition definition data each having a condition corresponding to input information, and defining transition from one process item in the plurality of process items to another process item;

an update step of accessing an externally located source of a new process item or transition definition data, acquiring the new process item or transition definition data, and updating an old process item stored at the process-item data storing step or old transition definition data stored at the transition-definition data storing step to the new process item or transition definition data;

the input information recognition and identification step of recognizing and identifying input information to be input; and

a step of selecting a piece of transition definition data based on comparison between the input information input and the conditions of the individual transition definition data, and transitioning a status to a process item designated by the selected transition definition data.

[0014] The present invention realizes a device control device, a speech recognition device, an agent device, and a device control method which can control a device adequately in response to an instruction uttered in the form of a language by a human being.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0015] Fig. 1 is a diagram showing a vehicular air conditioning system according to an embodiment of the present invention.

[0016] Fig. 2 is a more detailed structural diagram of the vehicular air conditioning system according to the embodiment.

- [0017] Fig. 3 is a diagram exemplarily showing a specific example of grouping flags.
- [0018] Fig. 4A and Fig 4B are diagrams for explaining a trigger acquisition process.
- [0019] Fig. 5A and Fig 5B are diagrams for explaining a discrimination process.
- [0020] Fig. 6A and Fig 6B are diagrams for explaining a discrimination process with an inquiry.
- [0021] Fig. 7A and Fig 7B are diagrams for explaining an output process.
- [0022] Fig. 8 is a diagram showing wires.
- [0023] Fig. 9 is a diagram showing a flow represented generally by a process item database and a wire database.
- [0024] Fig. 10A, Fig 10B, and Fig 10C are diagrams for explaining setting of a weighting factor.
- [0025] Fig. 11 is a flowchart for explaining the operation of an agent processing unit.
- [0026] Fig. 12 is a diagram showing the structure of an external server (distribution server) and a network structure.
- [0027] Fig. 13 is a flowchart for explaining a download process.

EXPLANATION OF REFERENCE NUMERALS

- [0028] 1 speech input unit
- 2 language analyzer
- 3 speech synthesis processing unit
- 4 speech output unit
- 5 input/output target device group
- 51 air conditioner
- 52 window opening/closing controller
- 6 agent processing unit

DETAILED DESCRIPTION OF THE INVENTION

[0029] An embodiment of the present invention will be explained below, taking a vehicular air conditioning system provided in a vehicle as an example, by referring to the accompanying drawings.

[0030] Fig. 1 is a block diagram showing the structure of the vehicular air conditioning system. FIG. 2 is a block diagram showing an example of the physical structures of individual section.

[0031] As shown in FIG. 1, the vehicular air conditioning system comprises a speech input unit 1, a language analyzer 2, a speech synthesis processing unit 3, a speech output unit 4, an input/output target device group 5, and an agent processing unit 6.

[0032] The speech input unit 1 receives a speech, generates speech data of a digital form from the input speech, and supplies the speech data to the language analyzer 2. Specifically, the speech input unit 1 comprises, for example, a microphone 11, an AF (Audio Frequency) amplifier 12, and an A/D (Analog-to-Digital) converter 13 incorporating a sample and hold circuit as shown in FIG. 2. The microphone 11 converts a speech to a speech signal, and outputs the signal. The AF amplifier 12 amplifies the speech signal from the microphone 11, and outputs the signal. The A/D converter 13 performs sampling and A/D conversion of the amplified speech signal from the AF amplifier 12, generating digital speech data, and supplies the data to the language analyzer 2.

[0033] As shown in FIG. 2, each of the language analyzer 2, the speech synthesis processing unit 3 and the agent processing unit 6 comprises a processor 21, 31, 61 comprising, for example, a CPU (Central Processing Unit), a non-volatile memory 22, 32, 62, such as a hard disk drive, which stores a program to be run by the processor 21, 31, 61, and a volatile memory 23, 33, 63, such as a RAM (Random Access Memory), which has a memory area to be a work area for the processor. A part or all of the functions of the language analyzer 2, the speech synthesis processing unit 3, and the agent processing unit 6 may be achieved by a single processor, or a single non-volatile memory and a single volatile memory.

[0034] The language analyzer 2 performs a speech recognition process on speech data supplied from the speech input unit 1. Through the speech recognition process, the language analyzer 2 specifies a candidate for a word represented by the speech data, and the likelihood (score) of the candidate. The scheme for speech recognition is arbitrary. A plurality of candidates may be specified for a word. The language analyzer 2 generates data indicating a specified candidate and the score of the candidate (hereinafter, called word data), and supplies them to the agent processing unit 6.

[0035] The non-volatile memory 32 of the speech synthesis processing unit 3 stores a speech segment database D1 which stores data representing waveforms of words, and a phoneme database D2 which stores data representing waveforms constituting phonemes.

[0036] The speech segment database D1 stores data representing waveforms of words. The phoneme database D2 stores data representing waveforms constituting phonemes. The speech synthesis processing unit 3 generates digital speech data representing a speech reading out text data supplied from the agent processing unit 6 by using the data stored in the speech segment database D1 and/or the phoneme database D2.

[0037] The speech synthesis processing unit 3 supplies the generated speech data to the speech output unit 4. The scheme of generating digital speech data is arbitrary; for example, slot filling method and rule-based synthesis (Rule-based synthesis) can be used. The slot filling method is a method by which, for example, word-by-word speeches are read by an announcer, and are linked together and output. The rule-based synthesis is a method by which relatively small units, such as phonemes (consonants or vowels) or kana, are linked together and output.

[0038] The speech output unit 4 reproduces a speech representing digital speech data supplied from the speech synthesis processing unit 3. More specifically, the speech output unit 4 has a D/A (Digital-to-Analog) converter 41, an AF amplifier 42, and a speaker 43 as shown in FIG. 2. The D/A converter 41 performs DA conversion on digital speech data to convert the data to an analog speech signal. The AF amplifier 42 amplifies the analog speech signal, and reproduces and outputs a speech represented by the analog speech data.

[0039] The input/output target device group 5 comprises, for example, an air conditioner 51, and a window opening/closing controller 52.

[0040] The air conditioner 51 performs a cooling, a heating or a fanning operation. The air conditioner 51 outputs data representing its operational status, e.g., status data indicating in which one of statuses, a cooling operation, a heating operation, temperature adjustment to a set temperature, a fanning operation, and at rest, the air conditioner is.

[0041] The window opening/closing controller 52, which opens or closes a window according to a control signal, includes a motor, a control circuit that controls rotation and stopping of the motor according to a control signal, and a winch that moves a window opening/closing control frame. The control circuit of the window opening/closing controller



52 outputs data representing, e.g., data indicating the amount (degree; the degree of opening) of the window open.

[0042] The agent processing unit 6 has a communication controller 64 comprising a modem and a packet communication terminal, and connects to a distribution server 100 to be  
5 discussed later over an external network (e.g., the Internet via a wireless telephone circuit) via the communication controller.

[0043] The non-volatile memory 62 of the agent processing unit 6 stores a word database D3. The word database D3 stores data of a plurality of words and one or more flags for word grouping indicating a plurality of word groupings in association with one another.

10 [0044] Individual flags associated with a single word are grouped in association with certain concepts. When a flag indicates a predetermined value (hereinafter, let the value be "1"), a word associated with the flag is grouped into a group associated with the flag. When the flag indicates another value (for example, "0"), the word is not grouped into the group associated with the flag.

15 [0045] FIG. 3 is a diagram exemplarily showing a specific example of grouping flags. In the example of FIG. 3, word grouping flags of four bits are associated with each of words "rise", "hot" and "open". The most significant bit (MSB) flag in the 4-bit bit group is associated with a concept "temperature". The second bit flag from the MSB is associated with a concept "operation of air conditioner". The third bit flag from the MSB is associated  
20 with a concept "opening/closing of window". The least significant bit flag is associated with a concept "failure".

[0046] As illustrated, the value of the 4-bit flags associated with the word "rise" is a binary number "1110", the value of the flags associated with the word "hot" is a binary number "1100", and the value of the flags associated with the word "open" is a binary number "1010".

25 [0047] Therefore, the flags indicate that the words "rise", "hot" and "open" are grouped under the concept "temperature", the words "rise" and "hot" are grouped under the concept "operation of air conditioner", the words "hot" and "open" are grouped under the concept "opening/closing of window", and none of the words "rise", "hot" and "open" are grouped under the concept "failure".

[0048] The individual words and the individual concepts are used as detection targets in a trigger detection process and discrimination targets for individual discrimination processes, which are stored in a process item database D4.

[0049] The non-volatile memory of the agent processing unit 6 further stores the process item database D4 and a wire database D5.

[0050] The process item database D4 is a database storing data (process item data) which describes the contents of various processes to be executed by the agent processing unit 6, e.g., a trigger acquisition process (TGxx), a discrimination process (Cnxx or QBxx), and an input/output process (EXxx to be discussed later), process item (pointer) by process item.

Note that "xx" is an identification number.

[0051] Of the process items stored in the process item database D4, data describing the contents of the "trigger acquisition process (TGxx)" includes trigger data (data designating the contents of data to be acquired as a trigger) which specifies a trigger to start those processes and a transition constant k for determining the progress direction (a constant which indicates the degree of transition constant in the progress direction and becomes a calculation standard for a weighting factor J to be described later).

[0052] Trigger data is arbitrary, and is, for example, data indicating in which one of the statuses, a cooling operation, a heating operation, temperature adjustment to a set temperature, a fanning operation, and at rest, the air conditioner is, data indicating the amount of a window open, data indicating the temperature in a room, or the aforementioned word data supplied from the language analyzer 2. Alternatively, trigger data may be data which is given from the process the agent processing unit 6 itself executes. When trigger data is (data which is acquired in the trigger acquisition process) is word data, it may be data indicating the "concept" that is assigned to the group to which a word represented by the word data belongs. It is to be noted that the contents of the trigger acquisition process are described in such a way that a plurality of trigger acquisition processes do not acquire word data representing the same word.

[0053] FIG. 4A shows an example of a trigger acquisition process TGxx. In this example, a trigger TG01 is a process of acquiring the word "hot" as a trigger (identifying the word "hot"), and the transition constant k for determining whether or not to proceed (transition) to a process following that process is 0.8. FIG. 4B illustrates a flowchart for a trigger acquisition process TG01. A trigger acquisition process TG02 is a process of acquiring the

word "open". A trigger acquisition process TG03 is a process of acquiring a word belonging to a group associated with the concept "temperature" (acquiring one of "rise", "hot" and "open" in FIG. 3). TG04 is a process of a word belonging to a group associated with the concept "operation of air conditioner" (acquiring either "rise" or "hot" in FIG. 3).

- 5    **[0054]** Data describing the contents of a "discrimination process (Cnxx)" stored in the process item database D4 includes a discrimination condition, a list of possible results as discrimination results, a transition constant  $k$  for determining the progress direction for each discrimination result, a transition constant  $k$  in a return direction to be discussed later, and data described for each discrimination process.
- 10   **[0055]** FIG. 5A shows an example of a discrimination process CNxx. In this example, a discrimination process CN01 is a "process of discriminating whether a window is open or not", and the transition constant  $k$  for determining whether or not to proceed to a following process when it is discriminated as open is 0.3, while the transition constant  $k$  for determining whether or not to proceed to a following process when it is not discriminated as open is 0.4.
- 15   FIG. 5B illustrates a flowchart for this example. A node CN01.1 shown in FIG. 5B is a start node indicating the start point of the process, a node CN01.2 is a node in the progress direction when it is discriminated that a window is closed, and the transition constant  $k$  is 0.4. Further, a node CN01.3 is a node in the progress direction when it is discriminated that a window is open, and the transition constant  $k$  is 0.3. A discrimination process CN02 is a
- 20   process of discriminating whether the status of the air conditioner 51 is on (in operation) or not, and the transition constant  $k$  for determining whether or not to proceed to a following process when it is discriminated as being in operation is 0.5, while the transition constant  $k$  for determining whether or not to proceed to a following process when it is not discriminated as being off (not operating) is 0.3.
- 25   **[0056]** In the "discrimination process", data to be used in discrimination may be acquired from an arbitrary acquisition source. Possible acquisition sources include, for example, other processes to be executed by the language analyzer 2 and the agent processing unit 6, devices belonging to the input/output target device group 5, and other external devices. In this case, data which describes the contents of the discrimination process may further include, for
- 30   example, data specifying an acquisition source for data to be used in discrimination.

**[0057]** In the "discrimination process", predetermined data may be output to a predetermined output destination prior to discrimination (in which case, a symbol indicating

the process is, for example, QBxx). For example, sending data representing a predetermined inquiry to the speech synthesis processing unit 3 prior to discrimination, or the like is possible. In case where predetermined data is output in the discrimination process prior to discrimination, data which describes the contents of the discrimination process includes, for example, the contents of data to be output and data designating the output destination of the data.

[0058] FIG. 6A shows an example of a discrimination process QBxx. In this example, a discrimination process QB01, for example, asks a user "Open a window? Activate the air conditioner?", and the transition constant  $k$  in the progress direction when the response (the user's answer) is "Activate the air conditioner" is 0.7, and the transition constant  $k$  in the progress direction when the response is "open a window" is 0.4. FIG. 6B illustrates a flowchart for this example. A node QB01.1 shown in FIG. 6B is a start node indicating the start point of the process, a node QB01.2 is a node in the progress direction when it is discriminated that activation of the "air conditioner" is designated in response to the inquiry, and the transition constant  $k$  is 0.7. Further, a node QB01.3 is a node in the progress direction when it is discriminated that opening a window is designated, and the transition constant  $k$  is 0.4. A discrimination process QB02 asks the user "Close a window?", the transition constant  $k$  in the progress direction when the response (the user's answer) is "Close" is 0.5, and the transition constant  $k$  in the progress direction when the response is "Do not close" is 0.3.

[0059] Data describing the contents of an "input/output process" stored in the process item database D4 comprises data which designates the contents of data to be input or output. Input data and output data may have arbitrary contents. For example, output data may be data representing reading of a speech which is generated by the speech output unit 4 via the speech synthesis processing unit 3 or a control signal which controls an external device. Input data may be, for example, data to be supplied from an external device.

[0060] FIG. 7A shows an example of an input/output process EXxx. In this example, an output process EX01, for example, is an operation to "Close a window, and activate the air conditioner", and the weighting factor  $k$  in the progress direction of executing a process after activation is 0.8. FIG. 7B illustrates a flowchart for this example. A node EX01.1 shown in FIG. 7B is a start node indicating the start point of the process, a node EX01.2 is a node indicating the end of the process, and the weighting factor  $k$  is 0.8. With regard to the

"input/output process", the weighting factor in the progress direction may not be set in which case selection of the node EX01.2 in the progress direction after execution of the "input/output process" should be made essential.

[0061] The wire database D5 comprises a set of data describing transitions among a plurality of processes (TG, CNxx, QBxx, EXxx) (the data will be hereinafter called wires). A wire comprises data described in, for example, a format illustrated in FIG. 8. As illustrated, a wire W<sub>n</sub> (W1, W2...) is data which designates a preceding process (X), a following process (Y), and a weighting factor J given to transition, for the transition (From (X) To (Y)) from the preceding process X (From (X)) to the following process Y (To(Y)). When the preceding process X is a discrimination process, it is necessary to describe transition from what discrimination result of the discrimination process. A transition causing process X and a transition target Y are specified by the node number of each process. The weighting factor J of each wire is not a fixed value, but is appropriately calculated and set according to the progress of a process. Calculation of the weighting factor J of the wire will be discussed later with reference to FIG. 10.

[0062] The agent processing unit 6 executes a flow generally represented by the process item database D4 and the wire database D5. The process item database D4 and the wire database D5 can describe a flow illustrated in FIG. 9 as a whole based on examples of FIGS. 3 to 8.

[0063] In the flow illustrated in FIG. 9, the agent processing unit 6 stands by for word data representing a word "hot" to be supplied from the language analyzer 2, and acquires the data when supplied, and passes it to a discrimination process CN01 (wire W1) in a trigger process TG01.

[0064] In the discrimination process CN01, the agent processing unit 6 acquires information representing whether or not a window is open from a window opening/closing controller 52, and a process proceeds to an input/output process EX01 when having discriminated that it is open (wire W2). In the input/output process EX01, a control signal to instruct closing of the window is output to the window opening/closing controller 52, and a control signal to instruct starting of the cooling operation is output to the air conditioner 51. As a result, the window opening/closing controller 52 closes the window, and the air conditioner 51 starts the cooling operation.

[0065] When it is discriminated in the discrimination process CN01 that the window is closed, a process proceeds to a discrimination process QB01 including an inquiry (wire W3). In the discrimination process QB01, first, the agent processing unit 6 supplies the speech synthesis processing unit 3 with data representing a text “Open a window? Or Turn on the air conditioner?”. The speech synthesis processing unit 3 reproduces a speech which reads out the text through the speech output unit 4.

[0066] In the discrimination process QB01, next, the agent processing unit 6 stands by for data representing a word “window” or words “air conditioner” to be supplied from the language analyzer 2, and when corresponding data is supplied, it is discriminated whether the data represents the word “window” or the words “air conditioner”. When it is discriminated that the data represents the word “window”, the process proceeds to an input/output process EX03 (wire W5), and when it is discriminated that the data represents the words “air conditioner”, the process proceeds to an input/output process EX02 (wire 06).

[0067] The agent processing unit 6 outputs a control signal instructing start of cooling to the air conditioner 51 in the input/output process EX02. The agent processing unit 6 outputs a control signal instructing opening of the window for ventilation to the window opening/closing controller 52 in the input/output process EX03.

[0068] The agent processing unit 6 stands by for data representing words “open a window” or “open” to be supplied from the language analyzer 2 (trigger process TG02), and shifts the process to the input/output process EX03 when it is supplied (wire W4).

[0069] In a case where the language analyzer 2 supplies plural pieces of word data, for example, the agent processing unit 6 executes a plurality of discrimination processes in parallel. In this case, if there are a plurality of processes involving the same word as an input target (for example, data input in a trigger acquisition process or a discrimination process) and word data representing the corresponding word is supplied from the language analyzer 2, the agent processing unit 6 executes all of those processes in parallel.

[0070] Next, a method of calculating the weighting factors J of the individual wires will be explained with reference to FIG. 10A to 10C.

[0071] In a case where a plurality of process are linked together by wires W, the weighting factor J of a wire W of interest is obtained by sequentially multiplying transition constants k

over a transition path originating from a process currently in process to the wire W of interest.

[0072] To facilitate understanding, a process whose flow is illustrated in FIG. 10A is assumed. In the general process in FIG. 10A, when a user utters “hot”, such is detected, and it is first discriminated whether or not a window is open, the user is asked whether to close the window or activate the air conditioner, and a process according to the user’s reply is executed.

[0073] In the process illustrated in FIG. 10A, for any of processes TG01, CN101, and QB101, the transition constants k for determining individual progress directions are all 0.5 in the forward direction, and are all 0.1 in the reverse (return) direction. In this case, wires are defined as illustrated in, for example, FIG. 10B.

[0074] For example, when a process (or control) is located at the trigger acquisition process TG101 (when a process pointer PP points the trigger acquisition process TG101), the agent processing unit 6 calculates the individual weighting factors J of wires W51 to W55 (all including non-illustrated wires if they are present) by sequentially multiplying transition constants k over a path with the trigger acquisition process TG101 up to a wire of interest as a starting point, writes calculation results in the wire database D5 in association with the wires W51 to W55.

[0075] Specifically, when a process is located at the trigger acquisition process TG101, the weighting factor J of the wire W51, for example, becomes the value of a transition constant k allocated to a node linked to the discrimination process CN101 of the trigger acquisition process TG101, i.e., 0.5. The weighting factor J of the wire W52 in the discrimination process CN101 becomes equal to a result of multiplying a transition constant  $k = 0.5$  associated with the wire W51 in the trigger acquisition process TG101 by a transition constant  $k = 0.5$  allocated to a node linked to the wire W52 in the discrimination process CN101, i.e., 0.25. Likewise, the weighing factor J of the wire W54 in the discrimination process CN101 becomes equal to, a resultant of multiplying a transition constant  $k = 0.5$  allocated to a node linked to the wire W51 in the trigger acquisition process TG101 by a transition constant  $k = 0.5$  of a node linked to the wire W54 in the discrimination process CN101, i.e., 0.25.

[0076] The weighing factor J of the wire W53 becomes equal to a result of further multiplying a result of multiplication of the transition constant  $k = 0.5$  allocated to the node

linked to the wire 51 in the trigger acquisition process TG101 by the transition constant  $k = 0.5$  allocated to the node linked to the wire W52 in the discrimination process CN101 by a transition constant  $k = 0.5$  allocated to a node linked to the wire W53 in the discrimination process QB101, i.e., 0.125. Likewise, the weighing factor J of the wire W55 becomes equal to a result of further multiplying a result multiplication of the transition constant  $k = 0.5$  allocated to the node linked to the wire W51 in the trigger acquisition process TG101 by the transition constant  $k = 0.5$  allocated to the node linked to the wire W52 in the trigger discrimination process CN101 by a transition constant  $k = 0.5$  allocated to a node linked to the wire W55 in the discrimination process QB101, i.e., 0.125.

[0077] Next, as the process transitions to the discrimination process CN101, the weighting factor J of the wire W52 becomes a value 0.5 equal to the transition constant  $k$  allocated to the node linked to the wire W52, and the weighting factor J of the wire 54 becomes a value 0.5 equal to the transition constant  $k$  allocated to the node linked to the wire W54. The weighing factor J of the wire W53 becomes equal to the product of the transition constant  $k = 0.5$  allocated to the node linked to the wire W52 and the transition constant  $k = 0.5$  allocated to the node linked to the wire W53 in the discrimination process QB101, i.e., 0.25, and the weighing factor J of the wire W55 becomes equal to the product of the transition constant  $k = 0.5$  allocated to the node linked to the wire W52 and the transition constant  $k = 0.5$  allocated to the node linked to the wire W55 in the discrimination process QB101, i.e., 0.25. Further, the weighting factor J in the reverse direction (direction of returning to the trigger acquisition process TG101) becomes a value 0.1 equal to a transition constant  $k = 0.1$  of the reverse-direction node allocated to the node linked to the wire W51.

[0078] Further, as the process transitions to the discrimination process QB101, both of the weighting factors J of the wires W53 and W55 become a value 0.5 equal to transition constants  $k$  respectively allocated to the nodes linked to the wires W53 and W55.

[0079] Further, the weighting factor J of the wire W52 becomes a reverse-direction transition constant  $k = 0.1$  allocated to a reverse-direction node linked to it. Further, the weighting factor J of the wire W51 becomes 0.01 that is the product of a reverse-direction transition constant  $k = 0.1$  allocated to a node linked to the wire W52 in the discrimination process QB101 and a reverse-direction transition constant  $k = 0.1$  allocated to a reverse-direction node linked to the wire W51 in the discrimination process CN101. The weighting factor J of the wire W54 becomes 0.05 that is a product of the reverse-direction transition



constant  $k = 0.1$  allocated to the node linked to the wire W52 in the discrimination process QB101 and a transition constant  $k = 0.5$  allocated to a node linked to the wire W54 in the discrimination process CN101.

[0080] Examples of changes in the weighting factors  $J$  of the individual wires  $W_n$  are  
5 illustrated in FIG. 10C.

[0081] Calculations of the weighting factors  $J$  are carried out not only for processes of the associated flow, but for all wires of the entire flows, and the calculated weighting factors  $J$  are set to the individual wires. A wire which is not associated with a current process may be allocated with a predetermined small coefficient. However, for a wire whose preceding  
10 process is a trigger acquisition process TGxx, the transition constant  $k$  is set high to some extent. This ensures a jump to a conversation whose contents greatly differ from the contents of a conversation that has been made just before.

[0082] Next, the operation of the entire system structured in this manner will be explained with reference to FIG. 11.

[0083] The speech input unit 1 and the language analyzer 2 individually operate, capture a  
15 speech, analyze it, and provide the agent processing unit 6 with word data. When (one or plural pieces of) word data associated with discrimination conditions are supplied from the language analyzer 2, the agent processing unit 6 executes the following processes.

[0084] First, the agent processing unit 6 recognizes (identifies) a supplied word (FIG. 11,  
20 step S11), and discriminates whether or not it corresponds to a word registered in the word database D4 (step S12). If it is not registered (step S12, No), a word inputting process is terminated.

[0085] If it is registered (step S12, Yes), on the other hand, the product of the likelihood (score)  $S$  of the word and the weighting factor  $J$  of the wire is calculated with respect to a  
25 process whose condition is the “concept” of the word or the group to which the word belongs (step S13).

[0086] For example, suppose that the process pointer PP indicates the trigger acquisition process TG101 in a case where the flow illustrated in FIG. 10A is being executed. The weighting factors  $J$  of the individual wires in this case are as illustrated in FIG. 10(c-1).

[0087] Suppose that in this status, word data indicating a word “hot” with a score of 80% and a word “window” with a score of 50% is input from the language analyzer 2.

[0088] In the example illustrated in FIG. 10A, the word “hot” relates to discrimination in the trigger acquisition process TG101, and the word “window” relates to discrimination in the discrimination process QB101.

[0089] As illustrated in FIG. 10(c-1), the weighting factor J of the wire W51 whose preceding process is a process of inputting the word data indicating the word “hot” is 0.5, and the weighting factor J of the wire W53 whose preceding process is a process of inputting the word data indicating the word “window” is 0.125. In this case, the products of the likelihoods S determined for the wires W51 and W53 and the weighting factors J are as indicated by equations 1 and 2.

[0090] (Equation 1) Product J·S of likelihood S and weighting factor J of wire W51: score of 80% for “hot” x weighting factor J (= 0.5) of wire W51 = 40

[0091] (Equation 2) Product J·S of likelihood S and weighting factor J of wire W53: score of 50% for “window” x weighting factor 0.125 of wire W53 = 6.25

[0092] The agent processing unit 6 carries out the above-described processes of determining the products of scores and weighting factors for all the wires the flow has.

[0093] Subsequently, the agent controlling unit 6 selects wire with the largest calculated product S·J (FIG. 11, step S14). The agent controlling unit 6 shifts control to a process following the selected wire (step S15). For example, when the product acquired for the wire W51 represents the most largest value, the input data is recognized as representing the word “hot”, and transition takes place to the discrimination process CN101 which is the process following the wire W51. Normally, with a process currently in process being taken as a starting point, the weighting factor J of the wire associated with the current process is relatively large. Accordingly, the process generally shifts to the next process, but in a case where a word which quite differs from a previous one and has a large likelihood is input, a process corresponding to that word may start.

[0094] When the process transitions, the agent processing unit 6 recalculates the weighting factors J of the individual wires based on the status after transition (step S16).

[0095] Afterward, the process is carried out in accordance with the contents of the process (step S17). In this example, the discrimination process CN101 is executed. That is, the agent processing unit 6 acquires information on opening/closing of a window from the window opening/closing controller 52, and carries out the process based on the acquired information.

5 For a signal indicating opening/closing of the window from the window opening/closing controller 52, the likelihood of opening may be 100% or 0%, the likelihood of closing may be 0% or 100%. They may be changed in accordance with the level of opening. When there is an input/output process EXxx during the process and the agent processing unit 6 executes this process and outputs speech data, it is generated as a speech.

10 [0096] When a process of controlling the air conditioner and window opening/closing controller 52, like the input/output process EX01, is executed during the process, the air conditioner 51 and the window opening/closing controller 52 are controlled in accordance with the contents of the process.

[0097] Although the discrimination process based on a “word” has mainly been explained  
15 referring to FIG. 8 and FIG. 9, the same is true of the discrimination process based on a concept. A wire from a trigger acquisition process based on a word or a concept is subject to discrimination. A somewhat large weighting factor may be set for transition from a trigger acquisition process. Specifically, a weighting factor 0.8 or the like is given to, for example, the wires in FIG. 8, i.e., transitions from the trigger acquisition process TG02 of acquiring  
20 word data representing words “open a window” in the flow in FIG. 9. Then, in a case where the process of the agent processing unit 6 is associated with a trigger acquisition process TG01, for example, as the user pronounces “open a window” and word data whose score for the words “open a window”, for example, is 90% is obtained, the product of the score and the weighting factor of the wire W5 associated with “determination on whether or not the user  
25 have said “Open a window”” is  $90\% \times 0.8$ , i.e., 72. If the value is larger than the products of scores and weighting factors J calculated for other wires, an input speech is recognized as “open a window”, and the probability that the process of the agent processing unit 6 jumps to the discrimination process QB101 becomes higher. If the weighting factors of the other wires are set extremely small, the probability of occurrence of transitions defined by the other wires  
30 becomes extremely low, resulting in that the control process is carried out with a recognition rate being improved along the flow of a conversation expected to some extent.

[0098] In the embodiment, transition in the reverse direction may occur. However, it is often undesirable to reverse a conversation in practice. Accordingly, a reverse-direction transition constant  $k$  should be set to a smaller value than a preceding-direction transition constant  $k$ . Then, even if speech data with a high score is obtained from an input speech, a product acquired for a wire written as a weighting factor  $J$  based on the reverse-direction transition constant  $k$  becomes a small value, so that the possibility of transition in the reverse direction can be suppressed at low. The agent processing unit 6 may treat a process in which the value of the acquired product does not meet a predetermined condition (a process in which a product value does not reach a predetermined value) in such a way that the process is excluded from transition execution targets.

[0099] As illustrated in, for example, FIG. 8, a wire defines transition in the form of from a process item to a process item. As a wire is described in the form as illustrated in FIG. 8 and stored in the wire database D5, it is possible to define relationships among the individual process items as if they were macro processes of a computer. This can facilitate connection of the individual process items.

[0100] Because a process item to be a trigger actually becomes discrimination of the score of a word or the like to be recognized associated with a wire to be connected, a trigger acquisition process item in the wire is not defined as the starting point of the wire, but the wire itself is defined as the origin of transition.

[0101] Further, as the connection relationships among the individual process items are defined by wires, wires can be easily added. In a case where the user frequently inputs a speech “find a family restaurant” with an intention to take a rest after a speech “hot” is input, for example, a wire is automatically added with respect to the search process item of a family restaurant. Then, after the wire is automatically added, it becomes possible to appropriately cope with the input “find a family restaurant” by increasing the weighting factor of a wire connected to a family restaurant search process item to some extent. In this case, however, the agent processing unit 6 shall store map data or the like including, for example, information representing the location of a family restaurant, or access external map data or the like. The automatic addition of a wire may be automatically carried out when the number of jumps from one process item to another is counted and reaches a predetermined number.

[0102] The agent processing unit 6 has a function of updating the contents of the process item database D4 and the wire database D5 to externally supplied new process items and

wires. Specifically, for example, as the contents of process item data and/or a wire stored in an external server is updated, the external server notifies the agent processing unit 6 over a network that the process item data and/or wire is updated. In response to the notification, the agent processing unit 6 accesses the server over the network, and downloads new process  
5 item data and/or a new wire. Old process item data and/or an old wire stored in the local process item database D4 and wire database D5 is updated to the downloaded new process item data and/or wire. The agent processing unit 6 may access the external server, and download a database if it has been updated.

**[0103]** An example of the structure of an external server will be illustrated in FIG. 12.

10 **[0104]** This external server 100, which distributes word data, process item data, and/or wire data to the agent processing units 6 of a plurality of vehicles through a communication network, comprises a controller 110, a communication controller 120, a word database 130, a process item database 140, a wire database 150, an input unit 160, and an output unit 170.

**[0105]** The controller 110 comprises, for example, a CPU (Central Processing Unit),  
15 controls the individual units of the distribution server 100, and realizes individual processes to be discussed later by running predetermined programs.

**[0106]** The communication controller 120 comprises, for example, a communication  
apparatus such as a modem or a router, and controls communications between the distribution server 100 and the agent processing unit 6 (communication controller) of the vehicle over the  
20 network.

**[0107]** The structure of the network is optional. For example, a private line, a public line network, a cable television (CATV) network, a wireless communication network, a wire broadcasting network or the like can be employed.

**[0108]** The word database 130 comprises, for example, a rewritable storage device like a  
25 hard disk drive, and stores word data distributed by the distribution server 100 together with version information (for example, time stamp).

**[0109]** The process item database 140 comprises, for example, a rewritable storage device like a hard disk drive, and stores process item data to be distributed, together with version information (for example, time stamp).

[0110] The wire database 150 comprises, for example, a rewritable storage device like a hard disk drive, and stores wire data to be distributed, together with version information (for example, time stamp).

[0111] The manager of the distribution server 100 properly manipulates the input unit 160, and updates information in the individual DBs 130 to 150.

[0112] Upon reception of an update notification from the distribution server 100 through the communication controller (CCU) 64, the agent processing unit 6 of each vehicle starts, for example, a process in FIG. 13A and establishes a session (step S21), downloads updated data, and updates local databases (step S22). Or, the agent processing unit 6 of each vehicle regularly or properly accesses the distribution server 100 through the communication controller (CCU) 64 and establishes a session (step S31), and acquires the version information (update date or the like) of each database (step S32). It compares the versions of both data with each other (step S33), downloads data newer than data locally stored, and sets the data in the database (step S34).

[0113] Employment of such a structure ensures appropriate update of data of all vehicle control systems which can use the distribution server 100 by merely updating information recorded in the distribution server 100.

[0114] It is desirable that the distribution server 100 should store data compressed and encrypted, and the agent processing unit 6 should download the data, expand and decrypt the data, and set the data in the database. This suppresses communication traffic, and prevents information leakage.

[0115] It is not limited to the case where as explained above, the agent processing unit 6 is notified that process item data and/or a wire is updated. For example, in a case where process item and/or a wire corresponding to the user's request is not stored, it may detect that, access the server over the network, and download new process item data and/or a new wire. Further, in a case where an input/output target device group 5 to be connected to the agent processing unit 6 is newly added or a new function is added, the agent processing unit 6 may automatically detect that, access the server over the network, and download the new process item data and/or wire in such a manner as described above.

[0116] The agent processing unit 6 has a function of automatically creating a new wire, and may download a new process item together with a program describing what relationship a

wire to be set has with respect to the downloaded new process item and an existing process item.

[0117] If data and a wire representing the contents of a process are described appropriately, the above-described vehicular air conditioning system need not completely specify a device to be controlled and the contents of control to be added, but can respond to a language spoken by the user, assume the request of the user from the language, appropriately determine what control should be applied to which device, and apply control according to the result of the discrimination to a device.

[0118] Because process item data and a wire are updated to new ones at any time, there is always room for changing how the agent processing unit 6 responds, so that the user is not likely to get bored with interaction with the vehicular air conditioning system.

[0119] The structure of the vehicular air conditioning system is not limited to the above-described one.

[0120] For example, devices belonging to the input/output target device group 5 should not necessarily be devices which bring about a result directly satisfying the user's demand, and may comprise, for example, a device (for example, a display apparatus like a liquid crystal display) that controls an external display apparatus or the like which outputs a message for prompting the user to take a specific action.

[0121] The word database may store not only data representing a word, but also data representing a phrase of plural words, as an element of the word database, or may store data representing a part of a word or a phoneme, as an element of the word database. A word or the like should not necessarily be grouped under a specific concept, and even in a case where grouping is done, data to be used for carrying out grouping may not take the form of a set of flags.

[0122] The agent processing unit 6 may spontaneously download new process item data and a new wire without waiting for notification from an external supplier, or may start downloading new process item data and a new wire in response to word data supplied from the language analyzer 2.

[0123] The agent processing unit 6 may change the transition constant  $k$  given to a node of each process in accordance with a predetermined rule based on the number of times a

transition represented by the wire has been executed, and rewrite the wire so that the weighting factor J becomes the value after the change.

[0124] Specifically, for example, the wire database D5 stores the number of executions of a transition represented by each wire, with respect to that wire. The agent processing unit 6  
5 rewrites the value of the number every time the transition is executed, thereby incrementing the value of the number by 1, and rewrites the transition constant k for each wire to a value which is proportional to, for example, the number stored for the wire. As a weighting factor J is calculated based on the changed transition constant k, the weighting factor J is changed based on the number of times the transition represented by the wire W has been executed.

10 [0125] The agent processing unit 6 may change data to be output in a discrimination process and an input/output process in accordance with data passed to those processes, data input through those processes, and other optional conditions.

[0126] The vehicular air conditioning system may have a display apparatus (for example, a liquid crystal display or the like) for outputting an image under the control of the agent  
15 processing unit 6, and the agent processing unit 6 may control the display apparatus in such a way that it displays a predetermined image for each process in the input/output process and the discrimination process.

[0127] The agent processing unit 6 may collectively acquire plural pieces of word data continuously spoken or the like in one input process and one discrimination process. The  
20 agent processing unit 6 may specify under what concept plural pieces of word data collectively acquired are grouped to the same group, and use a part of or all of the acquired word data for use in a process only if the specified concept matches with the predetermined concept. When a plurality of words are input, the likelihoods (scores) of words of the input speech stored in the word database D3 may be acquired, and it may be recognized that a  
25 common concept is input only if those words belong to the common concept. The score of the concept may be obtained at this time. For example, the result of multiplying scores obtained for individual words may be adopted as the score of the concept. The common concept and the score obtained for the common concept are to be output to each wire.

[0128] The agent processing unit 6 may comprise a plurality of data processing units (e.g.,  
30 computers or the like) which are connected to one another and bear a flow formed generally by various processes, such as a trigger acquisition process, a discrimination process, and an input/output process, and wires. In this case, it is sufficient that each data processing unit



which constitutes the speech synthesis processing unit 3 stores data representing that portion in the overall flow executable by the speech synthesis processing unit 3 which is probably executed by the unit, as elements of the process item database or the wire database. If data to be stored by each data processing unit is such data which macro-defines that process which is  
5 executed by the data processing unit, it is easy to cause a plurality of data processing units to perform distributed processing.

[0129] The vehicular air conditioning system may include the speech input unit 1, and the language analyzer 2 or the speech output unit 4.

[0130] The speech input unit 1 may have a recording medium drive unit (e.g., a floppy  
10 (registered trademark) disk drive, a CD-ROM drive or an MO drive) which reads a waveform signal from a recording medium (e.g., a floppy (registered trademark) disk, a CD (Compact Disc), an MO (Magneto-Optical Disk)) where data representing speeches is recorded, and supplies the signal to the language analyzer 2.

[0131] A device control device according to the invention whose embodiment has been  
15 explained above can be realized not only by an exclusive system but also an ordinary computer system.

[0132] For example, the vehicular air conditioning system that executes the above-described processes can be constructed by installing a program for executing the operations of the speech input unit 1, the language analyzer 2, the speech synthesis processing unit 3, the  
20 speech output unit 4 and the agent processing unit 6 into a personal computer connected to the input/output target device group 5 from a recording medium where the program is stored. The personal computer that executes the program executes flows shown in FIG. 9 and FIG. 11 as processes equivalent to the operation of the vehicular air conditioning system in FIG. 1.

[0133] The program that allows a personal computer to perform the functions of the  
25 vehicular air conditioning system may be uploaded to, for example, a bulletin board system (BBS) of a communication line, and distributed via the communication line, or a carrier may be modulated with a signal representing the program, the acquired modulated wave may be transmitted, and a device which receives the modulated wave demodulates the modulated wave to restore the program. Then, the above-described processes can be executed by  
30 invoking the program, and running the program like other application programs under the control of the OS.

**[0134]** When the OS bears a part of the processes, or the OS constitutes a part of a single constituting element of the invention, a program excluding that part may be stored in a recording medium. In this case, it is also assumed that according to the invention, a program for executing the individual functions or steps to be executed by the computer is stored in the

5 recording medium.